Infection, Genetics and Evolution 59 (2018) 38–47

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

Epidemiological investigation and analysis of the NS5B gene and protein variability of non-primate hepacivirus in several horse cohorts in Rio de Janeiro state, Brazil

Andreza Soriano Figueiredo[a,*], Elisabeth Lampe[b], Pedro Pereira Lira Furtado de Albuquerque[a], Flávia Löwen Levy Chalhoub[c], Ana Maria Bispo de Filippis[c], Livia Melo Villar[b], Oswaldo Gonçalves Cruz[d], Marcelo Alves Pinto[a], Jaqueline Mendes de Oliveira[a]

[a] Laboratory of Technological Development in Virology, Oswaldo Cruz Institute, Oswaldo Cruz Foundation, Av. Brasil 4365, Manguinhos, Rio de Janeiro, RJ 21040-360, Brazil
[b] Laboratory of Viral Hepatitis, Oswaldo Cruz Institute, Oswaldo Cruz Foundation, Av. Brasil 4365, Manguinhos, Rio de Janeiro, RJ 21040-360, Brazil
[c] Flavivirus Laboratory, Oswaldo Cruz Institute, Oswaldo Cruz Foundation, Av. Brasil 4365, Manguinhos, Rio de Janeiro, RJ 21040-360, Brazil
[d] Programme of Scientific Computation, Oswaldo Cruz Foundation, Av. Brasil 4365, Manguinhos, Rio de Janeiro, RJ 21040-360, Brazil

ARTICLE INFO

Keywords:
Equine hepacivirus
Horse
Prevalence
Risk factors
NS5B
Brazil

ABSTRACT

Among the hepacivirus species recently described, the non-primate hepacivirus/hepacivirus A found in horses and donkeys is closely related to the human hepatitis C virus (HCV). Therefore, the equine is an attractive surrogate large animal model for the study of HCV therapy, pathogenesis and prophylaxis. Despite global efforts, epidemiological and genetic studies have not elucidated the risk factors, virus distribution or genetic variability of the hepacivirus A, which are also important issues for the equine welfare. Little information about this background scenery is available in Brazil. The aims of this study were to investigate potential risk factors associated with hepacivirus A infection among different horse cohorts throughout the state of Rio de Janeiro and to evaluate the diversity of the viral NS5B gene and protein. Hepacivirus A RNA was detected in horse cohorts from all geographical mesoregions, independent of horse activity or breed investigated. Statewide prevalence ranged from 4.0% to 27.5%. Potential risk factors such as geographical location and age of female horses were significantly associated with the presence of virus RNA. Phylogenetic analysis revealed the circulation of subtype 2 in all mesoregions. NS5B gene sequences clustered according to geographical origin, while the NS5B fragments did not allow discriminant analysis. The predicted NS5B protein showed marked conservation, especially in the thumb domain. In conclusion, the higher frequency of hepacivirus A RNA detection in horses bred for reproduction purposes as well as in young females suggests a direct link between reproduction practices and the virus's spread. Additional studies are necessary to understand the distribution of this genetically conserved hepacivirus.

1. Introduction

The recently expanded genus *Hepacivirus*, family *Flaviviridae*, comprises 14 hepacivirus species with different animal hosts (*Hepacivirus A, B, D-N*), including the renamed human hepatitis C virus - HCV (*Hepacivirus C*) (Smith et al., 2016). HCV infection has worldwide distribution with prevalence of 2.5% (Petruzziello et al., 2016). Due to the chronic and progressive nature of the infection, HCV is a major cause of liver cirrhosis and cancer (Perz et al., 2006). Despite the recent advances in antiviral treatment (Zopf et al., 2016), the development of a vaccine is a challenge, prompting a continuous search for appropriate experimental models (Vercauteren et al., 2014).

Animal hepaciviruses have been found in dogs, horses, donkeys, bats, rodents, cows and Old World monkeys (Baechlein et al., 2015; Burbelo et al., 2012; Corman et al., 2015; Drexler et al., 2013; Kapoor et al., 2011; Lauck et al., 2013; Walter et al., 2017). The non-primate hepacivirus/hepacivirus A, found in horses and donkeys, is more closely genetically related to HCV than any other animal hepacivirus, according to findings to date (Pfaender et al., 2014). Other similarities have been demonstrated, such as liver tropism (Pfaender et al., 2015; Scheel et al., 2015), clearance of infection (Pfaender et al., 2017; Ramsay et al., 2015) as well as disease progression to chronicity

* Corresponding author.
E-mail address: andreza.figueiredo@ioc.fiocruz.br (A.S. Figueiredo).

https://doi.org/10.1016/j.meegid.2018.01.017
Received 3 December 2017; Received in revised form 10 January 2018; Accepted 22 January 2018

(Gather et al., 2016a). Thus, horses could be an interesting surrogate large animal model to study the evolution, immunity, persistent infection and pathogenesis of the virus.

The distribution of the hepacivirus A is global, with seroprevalence as high as 61.8% and circulating viral RNA as high as 35% (Burbelo et al., 2012; Elia et al., 2017; Figueiredo et al., 2015; Kim et al., 2017; Lu et al., 2016; Lyons et al., 2012; Matsuu et al., 2015; Pronost et al., 2016; Reichert et al., 2017; Tanaka et al., 2014). In Brazil, the hepacivirus A infection is highly prevalent, as demonstrated by the detection of viral RNA as high as 19.1% and 15.2%, respectively, in the breeding colony of a veterinary college and in racing horses (Figueiredo et al., 2015; Gemaque et al., 2014). However, no studies have elucidated the risk factors associated with infection.

Molecular epidemiology studies have been mainly based on NS3 and NS5B genes, showing the circulation of two subtypes (Elia et al., 2017; Figueiredo et al., 2015; Kim et al., 2017; Lu et al., 2016; Pronost et al., 2016), with an overall moderate divergence compared to HCV (Figueiredo et al., 2015; Pfaender et al., 2014). However, viral and host factors that might influence the geographic distribution and genomic variability are still elusive.

The aim of this molecular epidemiological study was to investigate the presence of the hepacivirus A in different horse cohorts throughout the state of Rio de Janeiro, to identify potential risk factors associated with infection, and to evaluate virus diversity by means of phylogenetic inferences and amino acid variability of the viral NS5B.

2. Methods

2.1. Study population and hepacivirus A screening

The total horse population was composed of 231 animals from 12 municipalities in the 6 geographical mesoregions of Rio de Janeiro state, Brazil (Fig. 1). The variables age, sex, breed, activity, geographical location and breeding system were recorded for each horse. Serum samples were obtained from whole blood collected from January 2015 to October 2016 and stored at −80 °C until use. All sample collections conformed to the guidelines for the care and use of animals (CEUA-IOC license 047/2015). Total RNA was isolated from 200 μL of serum using the High-Pure Viral Nucleic Acid Kit (Roche, Mannheim, Baden-Württemberg, Germany), eluted in a final volume of 50 μL and converted to cDNA using random hexamers with GoScript Reverse Transcriptase (Promega, Madison, WI, USA). The hepacivirus A screening was performed by nested RT-PCR directed to the NS5B region, as previously described (Lyons et al., 2012).

2.2. Liver enzymes

Serum levels of aspartate aminotransferase (AST) and gamma-glutamyl transferase (γ-GT) were measured in the sera of positive horses with an automatic analyzer (Labmax Progress, Labtest) according to the manufacturer's instructions.

2.3. Viral RNA quantification

Quantitative reverse transcriptase polymerase chain reaction (RT-qPCR) was used to determine viral loads of positive horses with the RotorGene Q (Qiagen). A Prime Time TaqMan Assay (Integrated DNA Technologies, Coraville, IA) directed to the viral 5' untranslated region (5'-UTR) was based on previously described oligonucleotides (Burbelo et al., 2012). The assay standard curve and limit of quantification (LOQ) were obtained by a 10-fold dilution series of a plasmid containing a 5'-UTR viral fragment.

2.4. DNA sequencing and phylogenetic and protein analysis

The viral NS5B gene was segmentally amplified using 6 degenerate primer pairs designed in this study (Supplementary material 1) based on the viral variability according to available sequences in the GenBank. The DNA sequences were determined in both directions with primers that overlapped two close regions using an ABI Prism BigDye Terminator version 3.1 cycle sequencing kit and an ABI Prism 310 genetic analyzer (Thermo Fisher Scientific). Multiple sequence alignments were generated covering nearly the full length of NS5B (1740 bp) in the MEGA 7 software (Kumar et al., 2016). All sequences were deposited in the GenBank (accession numbers MF441508-MF441520). The evolutionary analysis of the hepacivirus A, based on NS5B gene, circulating in Rio de Janeiro state was performed with maximum likelihood trees under the substitution model GTR + G + I in MEGA 7, with 1000 bootstrap replicates to evaluate the robustness of the trees and values of internal branches. The analysis based on NS5B fragment was performed by Bayesian inference using the Bayesian Markov chain Monte Carlo (MCMC) statistical framework, implemented in the BEAST v1.8.1 package (Drummond et al., 2012) in the substitution model GTR + G + I. Nucleotide (nt) and amino acid (aa) distance matrices were calculated based on p-distance either with MEGA 7 or Sequence Editor, Database and Analysis Platform, SSE Version 1.3 (Simmonds, 2012). Nucleotide sequences retrieved from the GenBank and the Los Alamos HCV Data Base are listed in Supplementary material 2. Predicted protein was analyzed with Weblogo version 2.8.2 (Crooks et al., 2004).

2.5. Statistical analyses

The infection rate of hepacivirus A was analyzed according to the following variables: age, sex, breed, activity, geographical location and breeding system. Concentration of liver enzymes and viral loads were compared through Anova or the Kruskal-Wallis test when appropriate, using the software GraphPad Prism 5, version 5.01 (San Diego, CA, United States). Significance was set at $P < 0.05$. Possible risk factors associated with virus infection were analyzed by univariate logistic regression conducted with the generalized linear model (GLM) and multivariate logistic regression conducted with the generalized additive model (GAM) using the R Project for Statistical Computing v.3.4.3. For all models, odds ratios (OR) and 95% confidence intervals (CI) were calculated.

3. Results

3.1. Hepacivirus A infection: descriptive analysis

The presence of hepacivirus A in Rio de Janeiro state was investigated by screening for virus RNA in horses from the six geographical mesoregions, as shown in Fig. 1. The 16 farms investigated were located in the rural area of 12 municipalities and were dedicated to raising animals for sport (57.6%) followed by reproduction (38.9%) and entertainment (3.5%). The horses' breeds were Brazilian Saddle Horse (BSH), or *Mangalarga Marchador* (60.6%), Campolina Horse (32.9%), Quarter Horse (6.1%) and one pony. Seventy percent were female and 29.9% were male, aged from 3 to 396 months old. The mean number of horses per farm was 112.5 (10–420) and the mean number of horses investigated per farm was 14.4 (2 − 23).

All six geographical mesoregions presented positive horses, with an overall prevalence of 13.4% (31/231). Positive horses were present in 62.5% (10/16) of farms and in 75.0% (9/12) of municipalities investigated. Table 1 presents possible risk factors associated with infection, the respective number and proportion of negative and positive horses, and the relative frequency. Viral detection varied from 4.0% (1/25) to 27.5% (11/40) across the state. Detailed information of the hepacivirus A-positive horses are provided in Table 2. The mesoregions were grouped (Fig. 2) according to the presence of the virus into most prevalent (group I: Northwest, North and Baixadas), with 71.0% positive animals, and least prevalent (group II: Central, Metropolitan and